Citation: Goode, N., Salmon, P.M., Taylor, N.Z., Lenné, M.G. & Finch, C.F. (In press). Lost in translation: the validity of a systemic accident analysis method embedded in an incident reporting software tool. Theoretical Issues in Ergonomics Science. Accepted 11th February 2016. DOI: 10.1080/1463922X.2016.1154230

Lost in translation: the validity of a systemic accident analysis method embedded in an incident reporting software tool SP-PS

Goode, N.[1*], Salmon, P.M. [1], Taylor, N.Z. [1], Lenné, M.G. [2] & Finch, C.F. [3]

[1]Centre for Human Factors and Sociotechnical Systems, Faculty of Arts and Business, School of Social Sciences, University of the Sunshine Coast, Australia

[2]Monash Injury Research Institute, Monash University, Australia

[3]Australian Centre for Research into Injury in Sport and its Prevention, Federation University Australia, Australia

*Corresponding author: Natassia Goode, PhD

Mail: University of the Sunshine Coast, Locked Bag 4, Maroochydore DC, QLD, 4558, Australia

Email: ngoode@usc.edu.au

Phone: +61 7 5456 5850

Fax: +61 7 5430 2859

Lost in translation: the validity of a systemic accident analysis method embedded in an incident reporting software tool

Abstract

Despite the proposed advantages of systems accident analysis (SAA) methods for understanding incident-causation, they have not been widely adopted by practitioners. The aim of this study was to evaluate the criterion-referenced validity of an SAA method embedded within an incident reporting software tool. Thirteen practitioners used the tool to collect and analyse incident data within their organisation. The incident data was then also analysed by researchers experienced in using the SAA method. Overall, there were low levels of agreement between participants and researchers regarding the identification and classification of factors and relationships. The findings indicate the systems thinking principles underpinning the SAA method may have been "lost in translation", in that participants often identified only one or two factors and showed a poor understanding of how to identify relationships between factors. The methodological developments required to ensure that practitioners can validly apply the SAA method are discussed.

Relevance to human factors/ergonomics theory

This paper makes two key contributions to the literature. First, it addresses whether SAA methods can be translated into usable and practical tools for practitioners. The findings highlight the challenges in embedding knowledge about SAA within a software tool. Second, the paper evaluates the validity of an SAA method when used by practitioners to analyse their own incident data. In contrast, the majority of studies evaluating the validity of accident analysis methods use incident reports selected by researchers. This study highlights the unique challenges that practitioners face when attempting to analyse their own incident data.

Keywords: systems thinking; validity; accident analysis

Lost in translation: the validity of a systemic accident analysis method embedded in an incident reporting software tool

Introduction

The systems approach now underpins the majority of accident analysis research (Salmon et al., in press; Underwood et al., 2013). This approach is based on the idea that safety in sociotechnical systems is impacted by decisions and actions made at all levels of that system. Therefore, accidents are caused by multiple interacting factors that go beyond the immediate context of the incident itself. Accidents and safety are described as emergent properties arising from the interactions of components within a system (Leveson, 2011; Rasmussen, 1997). This is in contrast to earlier theories of accident causation, which largely viewed accidents as a product of human errors and technical failures within the immediate context of the incident (Salmon et al., 2011). Researchers have demonstrated the applicability of systems accident analysis methods (SAA), and their advantages over non-systemic methods, in a wide range of safety critical domains including space exploration (Johnson & Muniz de Almeida, 2008), aviation (Branford, 2011), rail (Underwood & Waterson, 2014), public health (Cassano-Piche, Vicente, & Jamieson, 2009), disaster management (Salmon, Goode, Archer, et al., 2014) road freight transport (Salmon et al., 2013) and led outdoor activities (Salmon, Goode, Lenné, Finch, & Cassell, 2014; Salmon, Goode, Taylor, Lenne, Dallat & Finch, In Press). As the advantages of SAA methods have now been demonstrated by researchers, a key challenge is ensuring that this new approach is translated into practice to improve understanding of accident causation in organisations (Underwood & Waterson, 2013; Underwood et al., 2016).

There is evidence that the systems approach is currently not widely adopted in practice. Studies of investigation practices and manuals suggest they are still underpinned by earlier models of accident causation (Leveson, 2011; Lundberg, Rollenhagen, & Hollnagel, 2009). Analyses of investigation reports show that they largely focus on the immediate context of the accident (Newnam & Goode, 2015; Read, Salmon, & Lenné, 2013). Underwood and colleagues (Underwood & Waterson, 2013; Underwood, Waterson & Braithwaite, 2016) have identified a number of barriers preventing the adoption and usage of SAA methods by practitioners, including a lack of awareness, lack of training opportunities, accessibility and lack communication of information, usability, resource constraints, and questions around the reliability and validity of SAA methods.

The Understanding and Preventing Led Outdoor Accidents Data System (UPLOADS) software tool represents an attempt to address this research-practice gap within the led outdoor activity sector in Australia. The software tool was designed to help led outdoor activity practitioners collect and then analyse incident data using an SAA method developed specifically for the domain (Salmon et al., in press). Prior to the development of UPLOADS, the incident reporting systems used in this domain either did not support accident analysis (Goode, Finch, Cassell, Lenne, & Salmon, 2014), or were underpinned by Hale's (1984) Dynamics of Accidents Model. This model proposes that "Environmental" and "Human Factors" hazards combine linearly to create an accident potential (Hale, 1984 as cited in Curtis, 1995a; also known as the Accident Potential Model see Priest & Gass, 2005). For example, the New Zealand Mountain Safety Council have developed an incident reporting system for commercial, educational, not-for-profit, and informal groups conducting led outdoor activities. The reporting form presents checklists of contributing factors under the categories Activity Leader (e.g. inadequate physical condition, judgement error, inadequate supervision), Activity Participants (e.g. inadequate mental condition, improper motivation, inadequate training/experience), Equipment (e.g. no equipment, wrong equipment) and

Environment (e.g. adverse weather, inadequate visibility, terrain) (Hill, 2011). These categories limit analyses to the immediate context of the incident, and prevent practitioners from applying a systems approach to the accidents occurring in their organisation. As UPLOADS has now been implemented in a number of organisations (Salmon et al., In Press), the next step towards bridging the research-practice gap is evaluating whether the UPLOADS software tool achieves the purpose for which it was designed. That is, does the tool provide adequate support for practitioners in analysing their own incident data from a system perspective?

Validity testing is a critical but often overlooked part of human factors methods development (Stanton & Young, 1999, 2003). One aspect of testing the validity of a method involves evaluating whether the end users (i.e. practitioners) are able to generate analyses that are accurate compared to a criterion. Many criterion-referenced validity studies reported in the literature compare the outputs of a method against actual observations (e.g. errors predicted and errors observed; Stanton & Young, 2003). However, this methodology does not readily apply to the analysis of incident reports. In similar cases, 'gold standard' analyses produced by the system developers or an expert panel have been used as a criterion when objective standards are not available (e.g. Cornelissen et al., 2014; Lenne et al., 2008). This type of comparison reveals whether practitioners use a method as intended by the system developers and whether the outputs generated by practitioners are similar to the gold standard generated by experts. Both of which are essential steps in determining whether it achieves the purpose for which it was designed. The aim of this type of study is not to judge practitioner's performance, but rather to identify the areas for further methodological development (Stanton & Young, 2003).

A related issue is establishing the validity of SAA methods when practitioners analyse their own incident data; that is, real incident data reported within their own organisation. This represents a critical gap in the accident analysis literature, as the majority of accident analysis method validation studies use incident reports that are selected or created by researchers (Olsen, 2013). While this increases experimental control, these reports are likely to be much clearer and more detailed than the reports that practitioners collect on an everyday basis. This likely artificially increases the reported validity of the methods, and conceals some of the challenges associated with applying the methods in practice. To date, the UPLOADS accident analysis method has only been evaluated using artificial incident coding tasks (Goode, Salmon, Lenne, & Finch, 2014; Taylor, Goode, Salmon, Lenne, & Finch, 2015). Therefore, one of the purposes of this study is to examine the challenges faced by practitioners when applying an SAA method to their own incident data.

In summary, the aim of this study is to evaluate the criterion-referenced validity of an SAA method embedded within an incident reporting software tool. The study involved led outdoor activity practitioners using UPLOADS to collect and analyse incident data within their own organisation, which was then analysed by the system developers. Two aspects of the analyses were compared: 1) the identification of contributing factors and relationships between them from incident reports; and 2) the classification of contributing factors and relationships between them using the SAA. The overall purpose of this study was to examine the extent to which the tool supports the application the SAA method in practice, and to identify areas for further development.

Method

Design

The study was a prospective trial. It involved participants using the UPLOADS Software Tool to collect and analyse incident data within their organisation over a three month period (June to August 2014). The study was approved by the University of the Sunshine Coast Human Ethics Committee.

Participants

Organisations were invited to participate in the trial via outdoor education and recreation peak body and professional membership association newsletters. Interested organisations were asked to invite a senior staff member in a safety-related role to participate in the study. That person was responsible for implementing UPLOADS within the organisation. This involved: entering all incident reports; analysing and managing the data; and providing training to other staff on reporting incidents (hence forth referred to as the participants in this study). Forty-three organisations volunteered to take part in the study.

In total, participants from twenty-three organisations used the Software Tool and sent the data to the research team. Thirteen participants had analysed and coded reports. This represents a 53% response rate for using the incident reporting component of the software tool, and 30% response rate for using the SAA tools. Of the 13 participants who analysed and coded reports, 7 were male and 6 female, all aged between 24 and 64 years. All held a management role within the organisation, and 9 led activities as part of their role.

Materials

The UPLOADS system includes: the SAA method; paper-based incident report forms; software comprising data collection tools and systems accident analysis tools; and training material. These materials are described in the following sections.

SAA method

The SAA method (shown in Fig. 1) is based on Rasmussen's (1997) risk management framework and associated Accimap technique, and was developed specifically for led outdoor activity domain. The method consists of a framework for representing the system of factors involved in incidents and a taxonomy for classifying the factors, and relationships between them, involved in incidents. The framework describes the "led outdoor activity system" as a hierarchy across multiple levels: equipment, environment and meteorological conditions; decisions and actions of leaders, participants and other actors at the scene of the incident; supervisory and management decisions and actions; activity centre management, planning and budgeting, local area government, schools and parents, regulatory bodies and associations; and government department decisions and actions.

The taxonomy consists of two levels of categories. Level 1 describes the actors (e.g. activity participants, activity leaders, field managers, schools, parents etc.), artefacts (e.g. equipment) and activity context. Level 2 describes specific contributing factors relating to each of these components. The taxonomy was tested and refined in a series of previous studies (Goode, Salmon, Lenne, & Finch, 2014; Taylor, Goode, Salmon, Lenne, & Finch, 2015). It was then embedded in the systems accident analysis tools (described below) to categories the factors and relationships identified from incident reports.

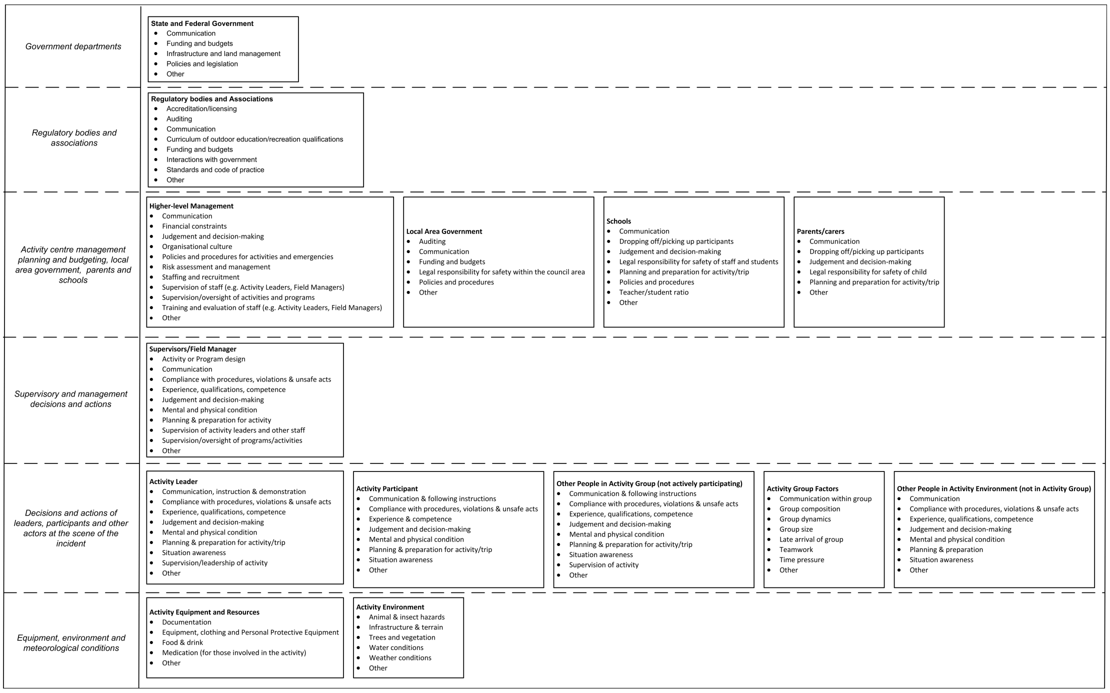


Figure 1 UPLOADS Accident Analysis Method, with framework and taxonomy, that is embedded within the software tool.

Paper-based incident report form

The paper-based incident report form comprises the fields presented Table 1. The form records both incidents associated with adverse outcomes and near misses. An "adverse outcome" is defined as any event resulting in a negative impact. A "near miss" is defined as a serious error or mishap that had the potential to cause an adverse outcome but failed to do so because of chance or because it is intercepted. The paper-based form includes an incident severity scale and a list of the factors shown in Fig. 1, so that reporters can reflect on the incident from a systems perspective. The list of factors is presented with the instructions “It is very important that you identify all the factors, and the relationships between them, which may have contributed to the incident you are reporting. To assist you in thinking about the causal factors involved in your incident, we have provided examples below of factors that have been found to play a role in previous incidents.”

The “reporter” in this context is the Activity Leader who was responsible for directly supervising the activity at the time of the incident. They complete all the relevant fields, and then pass the form onto their direct supervisor, who completes the field “Manager: explain in detail what your think caused the incident, including any relationships between causes, include suggestions, comments and recommendations.” The form is then passed on to the relevant person to enter the form into the Software.

Software

The software was developed in FileMaker Pro 12 and Java. The software consisted of: a data collection tool; a systems accident analysis tool for classifying the contributing factors and relationships identified from incident reports; an Accimap tool for summarising the data on contributing factor and relationships; and a tool for exporting de-identified data (e.g. names removed) to send to the research team. The data collection and systems accident analysis tool are described in the following sections.

Data Collection Tool. The data collection tool consisted of four linked databases (incidents; staff; clients; and participation). The "Incident" database captured the fields presented in the paper-based form, shown in Table 1. The "Staff" database captured demographic information about people employed by the organisations that were involved in incidents (e.g. age, gender, qualifications etc. This database was linked to the fields "Incident reporter name", "Staff responsible for supervision at the time of the incident" and "Person affected" within the Incident Database. The "Client" database captured demographic information about people not employed by the organisations that were involved in incidents, such as participants, teachers, volunteers and parents. This database was linked to the field "Person affected". Information could be entered into the "Staff" and "Client" database prior to reports of incidents, during the activity planning stage. Finally, the "Participation" database collected information on participation rates within the organisation during each month (i.e. types of activities, the number of participants undertaking each activity, and the number of days on which the activity had been conducted). This information was collected to provide context for the number of incident reports pertaining to each type of activity across the organisations.

Table 1 Fields presented in the paper-based incident report form and how they are captured in the "Incident" database of the Software

Section of the report	Fields	Type of field in the database
1. Incident characteristics	Incident reporter name	Drop-down menu linked to the Staff database
	Was the reporter present at the incident?	Drop-down options
	Date/Time	Drop-down options
	State/Territory	Drop-down options
	Staff responsible for supervision at the time of the incident	Drop-down menu linked to the Staff database
	Type of incident (adverse outcome/near miss)	Drop-down options
	Actual severity rating	Drop-down options
	Potential severity rating	Drop-down options
	Activity associated with the incident	Drop-down options
	Main goals associated with activity	Free text field
	Weather at the time of incident	Drop-down options
	Number of people involved in activity (participants, activity leaders, supervisors, volunteers)	Free text fields
	Location of incident	Free text field
	Did the activity leader have relevant qualifications?	Drop-down options
2. Adverse outcomes *if applicable*	Person affected	Drop-down menu linked to the Staff and Client databases
	Experience in activity associated with the incident	Drop-down options
	Was the incident fatal	Drop-down options
	Injury type	Drop-down options
	Injury location	Drop-down options
	Illness	Drop-down options

	Evacuation method	Drop-down options
	Hospitalisation required?	Drop-down options
	Was emergency services called?	Drop-down options
	Briefly describe the social/psychological impacts on the person described above.	Free text field
	Briefly describe any treatment at the scene of the incident	Free text field
3. Description	Describe the incident in detail, include: who was involved, what happened, when it happened, where it happened and any equipment involved.	Free text field
	Describe any relevant events leading up to incident	Free text field
	Describe why the incident was a near miss (if applicable)	Free text field
4. Contributing factors and relationships	Reporter: explain in detail what you think caused the incident, including any relationships between causes, include suggestions, comments and recommendations.	Free text field
	Manager: explain in detail what you think caused the incident, including any relationships between causes, include suggestions, comments and recommendations.	Free text field

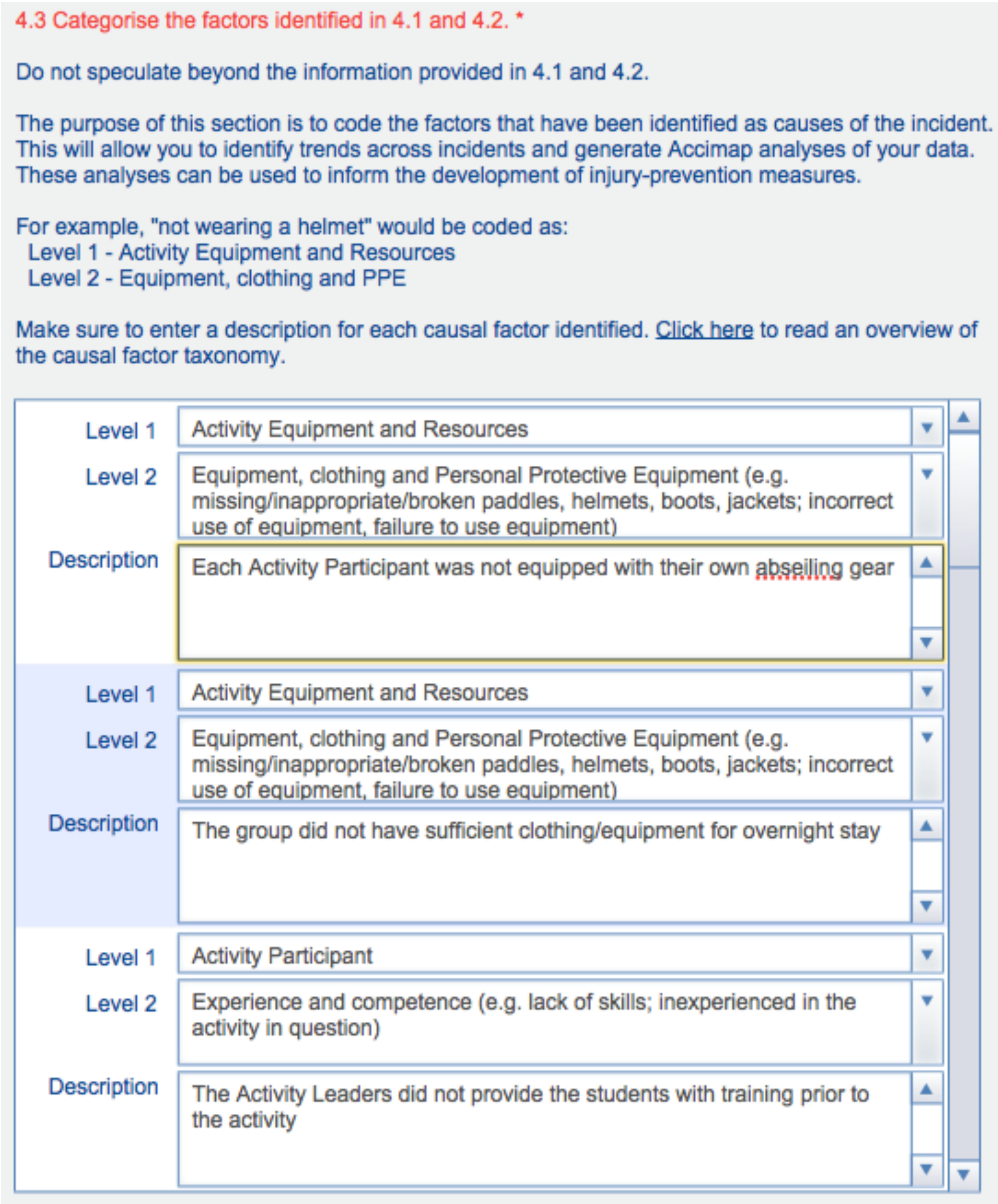
Systems Accident Analysis Tools. After each incident report has been entered into the "Incident" database, the program then prompts the user to code factors and relationships in the report.

The first section of the interface, shown in Figure 2, instructs the user to categorise the factors identified from the "Contributing factors and relationships" fields of the database (see Table 1). These fields describe the factors and relationships that the reporter and their manager feel contributed to the incident. Specific instructions are given not to speculate beyond the information provided in these fields of the report.

In Figure 2, the items presented in the drop-down menus labelled "Level 1" and "Level 2" are the corresponding codes from the SAA method taxonomy. In addition, the Level 2 drop-down menu presents examples of the specific factors that would be categorised under that code. For example, the level 2 code "Equipment, clothing and Personal Protective Equipment" is presented with the examples "missing/inappropriate/broken paddles, helmets, boots, jackets; incorrect use of equipment, failure to use equipment". There is also a link embedded into the interface that opens a comprehensive description of how to apply the SAA method (e.g. Click here to read an overview of the causal factor taxonomy).

For each factor they identify from the descriptions, the user first selects the Level 1 code from the drop down box that best describes that factor, and if possible, a corresponding Level 2 factor. The Level 2 drop-down menu only presents the factors that are relevant to the Level 1 code selected. The user can then enter a description of the factor (e.g. "Back pack rubbed clients back"). If they do not enter a description before they move onto the next factor, the software prompts them to do so with the message "Please enter description here" highlighted in red. The software manual instructs participants to copy and paste information from the "Contributing factors and relationships" fields into the description fields. An unlimited number of factors can be entered.

A similar procedure is followed for identifying and categorising relationships. The user interface is shown in Figure 3. For each relationship they identify from the descriptions, users are prompted to choose a pair of codes that best describes that relationship from the drop-down menus (e.g. “Higher-level Management: Training and evaluation of staff” and “Activity leader: Experience, qualifications, competence”). The drop-down menus only present the Level 2 codes that were used to describe factors in the previous task. The user can then enter a description of the relationship. An unlimited number of relationships can be entered.


4.3 Categorise the factors identified in 4.1 and 4.2. *

Do not speculate beyond the information provided in 4.1 and 4.2.

The purpose of this section is to code the factors that have been identified as causes of the incident. This will allow you to identify trends across incidents and generate Accimap analyses of your data. These analyses can be used to inform the development of injury-prevention measures.

For example, "not wearing a helmet" would be coded as:
Level 1 - Activity Equipment and Resources
Level 2 - Equipment, clothing and PPE

Make sure to enter a description for each causal factor identified. Click here to read an overview of the causal factor taxonomy.

Level 1: Activity Equipment and Resources
Level 2: Equipment, clothing and Personal Protective Equipment (e.g. missing/inappropriate/broken paddles, helmets, boots, jackets; incorrect use of equipment, failure to use equipment)
Description: Each Activity Participant was not equipped with their own abseiling gear

Level 1: Activity Equipment and Resources
Level 2: Equipment, clothing and Personal Protective Equipment (e.g. missing/inappropriate/broken paddles, helmets, boots, jackets; incorrect use of equipment, failure to use equipment)
Description: The group did not have sufficient clothing/equipment for overnight stay

Level 1: Activity Participant
Level 2: Experience and competence (e.g. lack of skills; inexperienced in the activity in question)
Description: The Activity Leaders did not provide the students with training prior to the activity

Figure 2 User interface for identifying and classifying the factors identified from the “Contributing factors and relationships” section of incident reports.

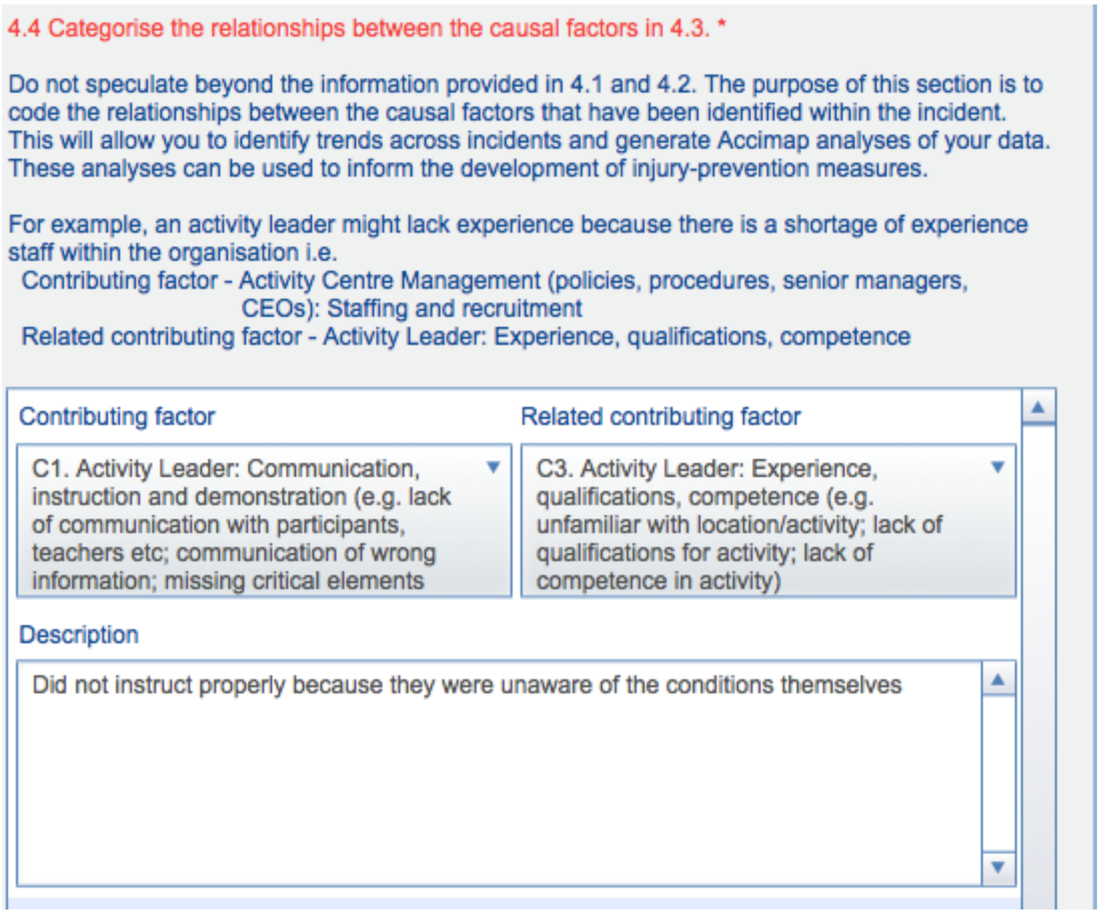


Figure 3 User interface for identifying and classifying the relationships identified from the "Contributing factors and relationships" section of incident reports.

Training material

Training material was developed for the person responsible for implementing UPLOADS in the organisation (the participant) and for staff members responsible for reporting incidents.

The training material for the person responsible for implementing UPLOADS was presented via a website. The website presented the following material: a manual explaining the accident causation theory underpinning the system, how to collect data about incidents and the SAA method; a manual describing how to use the software; and a series of videos demonstrating how to perform each task associated with the software (e.g. entering incidents reports, categorising factors and relationships, generating Accimaps). The manual describing how to use the software was an online document, with the videos embedded at the relevant points. In addition, links to relevant sections of the website (e.g. definitions, SAA method) were embedded in the Software.

The training material for staff members responsible for reporting incidents consisted of a PowerPoint presentation explaining: what types of incidents to report; the details required for the incident report forms; and the underpinning theory. In addition, the incident report form (described in a preceding section) was used to reinforce these instructions.

Procedure

On contacting the research team, organisations and participants were asked to provide written consent to participate in the study. Participants were then sent a link to a demographics questionnaire presented on Survey Monkey. Once completed, participants were sent an email with instructions describing their responsibilities in participating in the trial, which included: training other staff on how to report incidents from a systems perspective and complete reports, collecting reports, and using the Software. The initial instructions included details on: the type of incidents that should be reported; how to download the Software and supporting material; a link to the training materials website; and dates for submitting data to the research team.

The initial email also included instructions about utilising the training materials. Participants were instructed to first read through the manual describing the underpinning theory and SAA method. Second, read through the manual describing how to use the software and the videos demonstrating the tasks. Third, arrange training sessions for staff on how to report incidents, using the PowerPoint presentation provided. In addition, the email invited participants to contact the research team via phone or email if they had any questions or required any help.

At the end of the three-month trial period, participants were sent an email asking them to submit a deidentified version of the data they had collected. The software tool included a function that automatically removed all identifying information from the database to send to the research team.

Data analysis

The software tool explicitly instructs users to only identify and classify the factors and relationships described in the "Contributing factors and relationships" section of the incident reports they collect, and not to speculate beyond this information. Therefore, the analyses focussed on comparing participant's and researcher's identification and classification of factors from this section of the reports.

All data collected was merged into a central database. The set of incident reports that had been analysed by participants was then extracted. Each of these reports was associated with a list of contributing factors and relationships that had been identified and classified by participants. The list of factors was a table with the columns: "Level 1 code", "Level 2 code", and "Description of the factor". The list of relationships was a table with the columns: "Factor 1", "Factor 2", and "Description of the relationships". The "descriptions" are the factors and relationships that participants had identified from the "Contributing factors and relationships" sections of the reports (see Table 1), while the other columns are the codes from the taxonomy that had been selected to classify the causal factors and relationships.

The data were analysed by the 2 of the authors. One researcher has extensive experience in coding qualitative data including accident and incident reports. Both researchers have extensive experience in using the UPLOADS accident analysis method to analyse incident data. Importantly, both researchers (and the research team) also have extensive experience in both the led outdoor activity sector and the analysis of led outdoor activity incidents and are regularly exposed to incident data and incident analyses (e.g. Salmon et al., 2014).

First, the researchers identified the factors and relationships from the "Contributing factors and relationships" section of each incident report (see Table 1), and then discussed any discrepancies and reached a consensus for each description. This involved little subjective judgement on behalf of the researchers, as they only identified the factors and relationships that were explicitly stated. Second, they used the UPLOADS taxonomy to classify all the factors and relationships identified by both participants and themselves, to allow for accurate comparisons between the descriptions. The researchers then discussed any discrepancies between their classifications and reached a consensus for each code; there was 89% agreement between the two researchers on the selection of codes used to classify the factors identified by participants. For example, in one incident report, the "Contributing factors and relationships" section stated "Reporter explanation: pack was rubbing on her back. Padding and tape was applied to sore spot" and "Manager explanation: The bag was rubbing on [participants] back. [Participant] was carrying an ankle injury." The participant had identified the factor "Back pack rubbed clients back", which was classified by the researchers as "Activity equipment and resources: Equipment, clothing and Personal Protective Equipment". The researchers identified the factors "The bag was rubbing on participants back" and "[Participant] was carrying an ankle injury", which were classified as "Activity equipment and resources: Equipment, clothing and Personal Protective Equipment" and "Activity participant: mental and physical condition."

To examine whether participants and researchers identified the same factors and relationships from the "Contributing factors and relationships" section of each report, the agreement between the factors and relationships identified by participants and researchers was calculated. A "hit" was scored if participants identified the same factor or relationship from the report as researchers. A "miss" was scored if participants did not identify a factor or relationship that was identified by researchers. A "false alarm" was scored if a participant identified a factor that was not identified by researchers. The percentage of hits, misses, and false alarms for each participant across all incidents was then calculated.

To examine whether there were differences in the types of contributing factors identified by participants and researchers, the number of times each code from the taxonomy was used by participants and researchers across all reports was calculated. The frequencies for each code were then represented on the framework presented in Figure 1.

To examine whether participants and researchers used the same codes to classify factors and relationships, the codes used to classify the descriptions of factors and relationships that had been identified by participants from each incident report were compared. A "hit" was scored if participants selected the same code/s from the taxonomy as the researchers to classify a description of a particular factor or relationship. For example, for the description "Slippery river banks" a hit would be scored if researchers and participants both selected "Activity Environment: Infrastructure and terrain". A "miss" was scored if participants chose a different code to the researchers. A "false alarm" was scored if participants selected a code from the taxonomy but did not provide a description or provided too little information to support a classification. The percentage of hits, misses, and false alarms for each participant was then calculated.

To examine whether there were differences in the types of codes selected by participants and researchers to classify the factors identified by participants, the total number of hits and misses across all participants for each code in the taxonomy was calculated. The total number of hits and the total number of misses for each code were then represented on the framework presented in Figure 1.

Results

Use of the SAA Tools

Of the 13 participants who used the SAA tools, 149 incident reports had been analysed out of a total of 228 reports. Table 2 presents a summary of the extent to which each of the participants used the SAA Tools within the software. One participant did not identify or classify any relationships, and one participant entered codes to describe relationships without corresponding descriptions.

The median numbers of factors and relationships entered into the SAA Tool by participants per report was 2 (range 1 to 5, n = 149 reports) and 1 (range 0 to 6, n = 80 reports), respectively. The median numbers of factors and relationships identified by researchers from the "Contributing factors and relationships sections of the report" were 2 factors (range 0 to 6) and 1 relationship (range 0 to 4) per report.

Table 2 Summary of the extent to which each participant used the SAA tools within the software

Participant	Total # incidents entered	% of incidents with factors identified or classified	% incidents with relationships identified or classified
1	21	100	100
2	11	91	91
3	7	86	86
4	24	100	100
5	6	100	83
6	11	100	36
7	2	100	100
8	9	100	89
9	1	100	0
10	2	100	100
11	114	34	15
12	8	100	67
13	12	75	100

Comparison of identification of factors and relationships from incident reports

A summary of the agreement between the factors identified by participants and researchers from the incident reports is presented in Table 3. On average, 43% (SD = 20%, range = 0% to 73%) of the factors identified by participants were the same as those identified by researchers (hits); 57% (SD = 20%, range = 27% to 100%) of the factors identified by researchers were not identified by participants (misses); and 59% (SD = 20%, range = 35% to 100%) of the factors identified by participants were not identified by researchers (false alarms).

By examining the factors that were extracted from reports by participants and researchers, three issues were identified that could explain some of the differences. First, some participants systematically included factors that were evident from the other sections of the report, but not from the "Contributing factors and relationships" sections of the reports, contributing to a high rate of false alarms. For example, one report stated in the description of the incident: “Playing camouflage a hide and seek game student was running to find a hiding place. Dived onto a rock and cut knee.” The contributing factors and relationship section stated: “Student was over excited and competitive in the game and was not paying attention to surroundings.” Only the contributing factor “rocks” was identified from the description section of the report.

Second, some participants ignored factors that were explicitly stated in the “Contributing factors and relationships” section, resulting in a high rate of misses. In the previous example, the factor “not paying attention to surroundings” was not identified from the description of contributing factors provided by the reporter.

Third, some participants went beyond the information provided in the report altogether, contributing to the high rate of false alarms. The added information often related to failures relating to the Activity Participant. For example, from the information “Description: slipped over at the staircase mild abrasion in the left lower leg applied sterile dressing at the scene redressed the following morning” and “Contributing factors and relationships: only the child was involved with the incident”, the participant identified the factor “child did not take into account how the ground being wet would affect the way they walk”. In another example, the following information was provided “Description: Playing at free play time. Supervised. [Participant] received a small burn behind the left knee” and “Contributing factors and relationships: Rope and free play…rope being grabbed and hitting [participant].” The factors identified were “free play” and “inexperience with rope”; however, participant’s level of experience was not explicitly stated in the report.

Table 3 Summary of the agreement between the contributing factors identified by participants and researchers

	No of factors identified from reports		Hits		Misses		False alarms	
Participant	Researchers	Participants	n	%	n	%	n	%
1	21	25	8	38	13	62	17	68
2	28	25	10	36	18	64	15	60
3	26	20	9	35	17	65	11	55
4	41	46	30	73	11	27	16	35
5	13	15	9	69	4	31	6	40
6	17	11	7	41	10	59	4	36
7	20	18	9	45	11	55	9	50
8	5	4	2	40	3	60	2	50
9	5	17	3	60	2	40	14	82
10	2	1	0	0	2	100	1	100
11	7	7	1	14	6	86	6	86
12	70	88	40	57	30	43	48	55
13	16	16	8	50	8	50	8	50

A summary of the agreement between the relationships identified by participants and researchers from the reports is presented in Table 4. On average, 25% (SD = 29%, range = 0% to 100%) of the relationships identified by participants were the same as those identified by researchers; 75% (SD = 29%, 0 to 100%) of the relationships identified by researchers were not identified by participants; and 85% (SD = 14%, 3% to 100%) of the relationships identified by participants were not identified by researchers.

An examination of the relationships identified by participants revealed that they often tended to ignore relationships explicitly described in the “Contributing factors and relationships" section, resulting in a high rate of misses. For example, in one incident, the "Contributing factors and relationships section stated "Reporter: Mountain biking, first major hill all students didn't give enough space and a few people crashed" and "Manager: With many international students with minimal experience mountain biking, the briefing could have been more involved." No relationships were identified from this incident.

In addition, many of the relationships entered into the SAA tool by participants described only a single factor, rather than a relationship, resulting in the high rate of false alarms. For example, "Emotional stress", "Client tripped", "Participants over use of leg led to strain" and "Slippery banks."

Table 4: Summary of the agreement between the relationships identified by participants and researchers

	No. of relationships identified from reports		Hits		Misses		False alarms	
Participants	Researchers	Participant	n	%	n	%	n	%
1	5	8	1	20	4	80	7	88
2	12	23	2	17	10	83	21	91
3	14	10	3	21	11	79	7	70
4	23	28	11	48	12	52	17	61
5	7	9	3	43	4	57	6	67
6	7	5	0	0	7	100	5	100
7	-	-	-	-	-	-	-	-
8	1	1	0	0	1	100	1	100
9	1	6	1	100	0	0	5	83
10	-	-	-	-	-	-	-	-
11	4	3	0	0	4	100	3	100
12	19	24	3	16	16	84	21	88
13	7	6	1	14	6	86	5	83

Comparison of the types of factors identified by participants and researchers

Figure 4 presents a comparison of the types of factors identified by participants and researchers from the "Contributing factors and relationships" section of all incident reports (n = 141 reports). As Figure 4 shows, neither group identified factors at the "Government departments" nor "Regulatory bodies and associations" levels of the framework. Researchers identified more factors relating to "Higher-level management", "Supervisors/Field Managers", "Activity Leader" and "Activity Group Factors" than participants. Participants identified more factors relating to the "Activity Participant", specifically for codes relating to "Communication and following instruction", "Compliance with procedures, violations and unsafe acts." They also identified more factors relating to "Equipment, clothing and Personal Protective Equipment" and "Infrastructure and terrain". However, overall these differences were small.

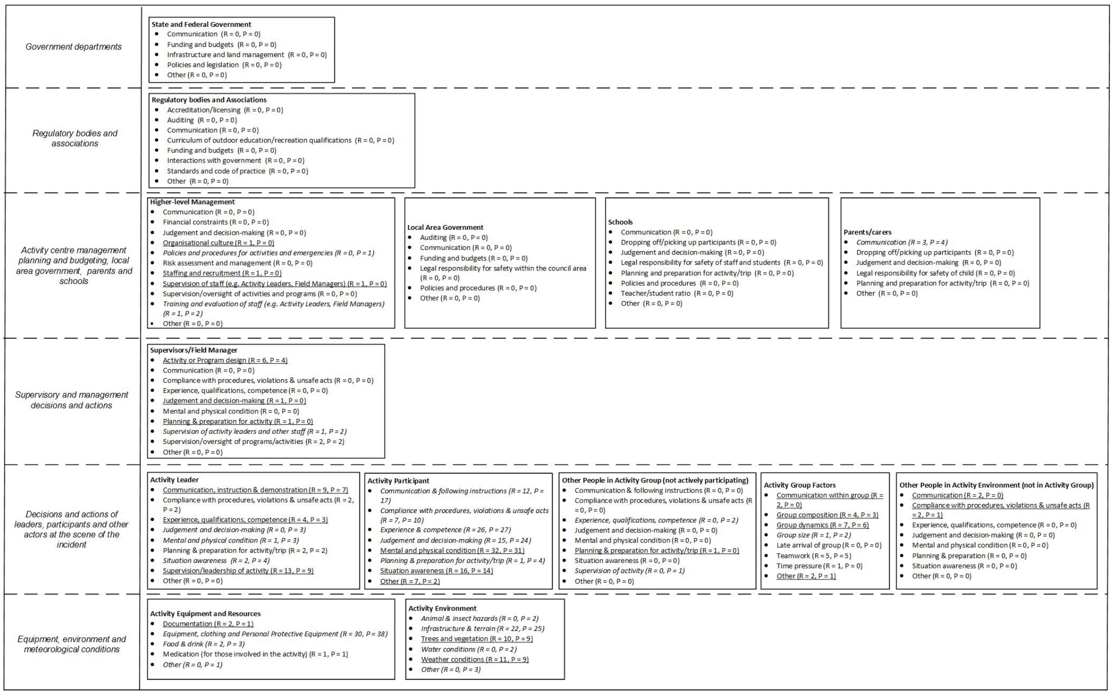


Figure 4 Comparison of the types of factors identified by participants and researchers from the "Contributing factors and relationships" section of all incident reports (n = 141 reports). Numbers in brackets indicate the frequency with which participants ("P") and researchers ("R") identified each type of factor. The factors identified more frequently by researchers are underlined. The factors identified more frequently by participants are italicised.

Comparison of classification of factors and relationships by participants and researchers

A summary of the agreement between the codes selected by participants and researchers to classify the factors identified by participants is presented in Table 5. On average, 56% (SD = 23%) of the codes selected by participants were the same as the researchers; 37% (SD = 23%) of codes selected by participants were not the same as those selected by researchers; and 7% (SD = 18%) of codes selected by participants had insufficient information in the description to support a selection by the researchers.

Table 5 Summary of the agreement between codes selected by participants and researchers to classify the contributing factors identified by participants

	Hits		Misses		False alarms	
Participant	n	%	n	%	n	%
1	12	48	12	48	1	4
2	15	60	10	40	0	0
3	15	75	5	25	0	0
4	44	86	7	14	0	0
5	11	69	5	31	0	0
6	9	30	1	3	20	67
7	2	50	2	50	0	0
8	12	60	7	35	1	5
9	0	0	1	100	0	0

10	4	50	4	50	0	0
11	65	71	27	29	0	0
12	13	76	4	24	0	0
13	12	48	9	36	4	16

A summary of the agreement between the codes selected by participants and researchers to classify relationships identified by participants is presented in Table 6. On average, 35% (SD = 25%) of the codes selected by participants were the same as the researchers, 45% (SD = 28%) of the codes selected by participants were not the same as the researchers, and 28% (SD = 33%) of codes selected by participants had insufficient information in the description to support a selection by the researchers.

An examination of the codes selected by participants revealed two consistent problems. First, in many cases participants inappropriately selected the same factor twice to describe a relationship. For example, to code the description "Participant ignored instructions and lacked supervision" a participant selected "Activity Participant: Communication and following instructions" for both factors, despite the fact they had identified "Activity Leader: Communication, instruction & demonstration" as a factor in previous task. This may indicate that there is a problem with the user interface shown in Figure 3. Second, there was often insufficient evidence to support a classification of a relationship. As described in the previous section, often only a single factor was described. In addition, for some relationships no description was provided to support the selection of the codes.

Table 6 Summary of the agreement between codes selected by participants and researchers to classify the relationships identified by participants

	Hits		Misses		False alarms	
Participant	n	%	n	%	n	%
1	0	0	7	30	16	70
2	2	8	21	88	1	4
3	7	64	3	27	1	9
4	19	66	9	31	1	3
5	5	56	4	44	0	0
6	2	33	3	50	1	17
7	1	50	0	0	1	50
8	2	22	2	22	5	56
9	2	67	1	33	0	0
10	-	-	-	-	-	-
11	8	31	15	58	3	12
12	2	29	4	57	1	14
13	0	0	12	100	12	100

Comparison of the types of codes used to classify factors

Figure 5 present a summary of the total number of hits and total number of misses across all participants for each code in the taxonomy. From Figure 5, two consistent areas for disagreement can be identified. First, 23 out of the 41 factors classified by researchers as

“Equipment, clothing and Personal Protective Equipment” were classified using other codes by participants. The descriptions of the factors show that most of these relate to the incorrect use of equipment. Second, most of the factors that researchers classified under “Activity Group” codes were classified under “Activity Participant” codes by researchers.

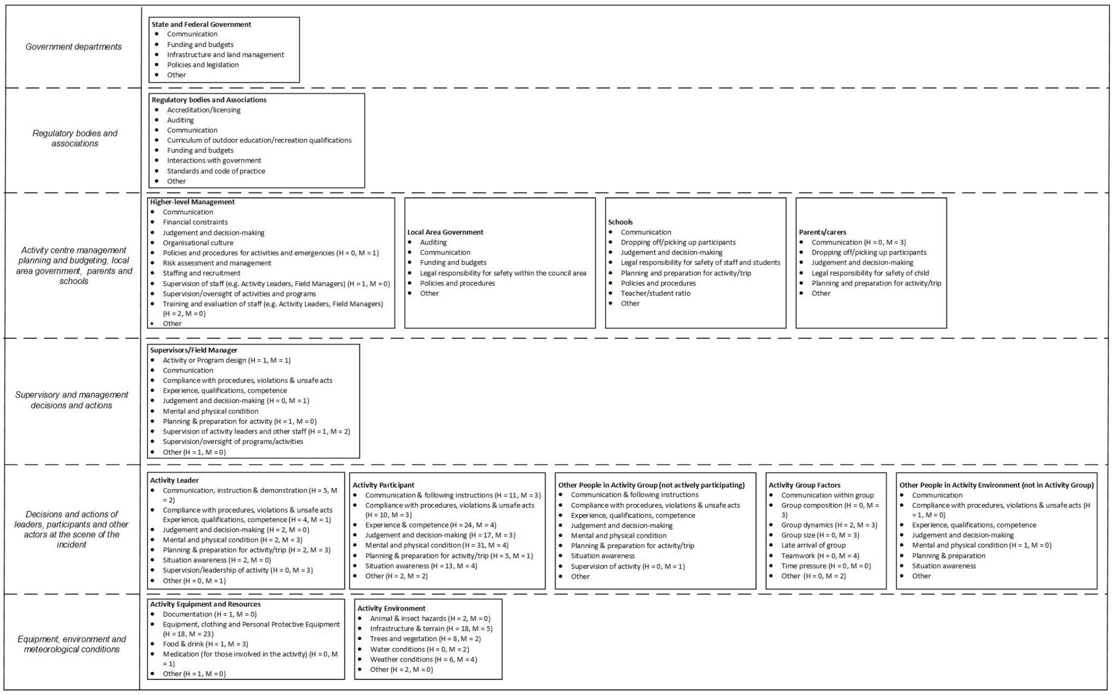


Figure 5 Summary of the total number of hits and total number of misses across all participants for each code in the taxonomy. Numbers in brackets indicate the total number of hits (“H”) and misses (“M”) associated with each code.

Discussion

The aim of this study was to evaluate the criterion-referenced validity of an SAA method embedded within an incident reporting software tool. The study examined how led outdoor activity practitioners used the software to identify and classify contributing factors and relationships from their own incident reports, compared to researchers experienced in both analysing led outdoor activity incident reports and using the method. On average, the level of agreement between participants and researchers regarding the identification and classification of factors was poor. However, participant’s scores were highly variable, with a few participants showing almost no agreement with researchers and a few participants close to what is considered an “acceptable level” (e.g. at least over 75%; Olsen, 2013). In contrast, the level of agreement between participants and researchers regarding the identification and classification of relationships was extremely poor across all participants. Overall, the findings indicate the principles of systems thinking may have been “lost in translation” through the process of implementing the SAA method within the software tool and application by practitioners.

While these findings are disappointing, they represent an important step in the development of a useable and valid SAA software tool for practitioners, and an important question is specifically what issues underpinned the low validity scores.

In terms of the initial identification of factors and relationships from incident reports, two key issues appear to be pertinent. First, it is clear that the instructions not to speculate beyond the information provided in the "Contributing factors and relationships" section of the incident reports unnecessarily constrained the analyses of the incidents. The instructions assume that the staff who complete the incident report (i.e. Activity Leaders and their direct managers or supervisors) will synthesise all the relevant information from the description and other fields into a list of contributing factors and relationships, which can then be entered into the SAA Tool. However, the findings suggest that this is not the case, as some participants consistently use additional information from the other fields in the report in their analyses of contributing factors. The additional factors they identified appeared reasonable, although researchers analyses were constrained to the information presented in the "Contributing factors and relationships" section of the incident reports. Changing the instructions within the software tool, and including examples in the training material on how to use all the information contained within the report to conduct the analyses, could potentially ameliorate this source of disagreement between participants and researchers.

A second issue is whether it is reasonable for the person undertaking the analysis to identify additional factors and relationships beyond the information provided in the report or exclude factors from the analysis that were explicitly stated in the "Contributing factors and relationships" section of the report. On the one hand, all the participants in the study were managers within the organisation where the incident had occurred, and were responsible for implementing the incident reporting system. Therefore, they are likely to have relevant background information that they can draw upon for incident analyses, particularly regarding organisational factors that may have contributed to the incident (e.g. a lack of funding impacting training programs). In addition, they are potentially in a position where they can make judgements regarding whether a factor contributed to an incident or not. On the other hand, it was often unclear why participants chose to include or exclude certain factors from their analyses. The additional factors identified by participants were often concerned with "blaming" the Activity Participant who had sustained the injury (e.g. poor decision-making, inexperience); it was unclear whether these inferences were drawn from the participant's experiences of the actual event or through speculation. In addition, participants often limited the analysis to only one or two factors within the immediate context of the incident, while excluding other factors that were explicitly stated in the "Contributing factors and relationships" section of the reports that appeared reasonable. Both of these issues potentially indicate a lack of understanding of the underpinning systems model of accident causation.

This conclusion is also supported by participant's poor analyses of relationships between factors. Participants often tended to ignore relationships explicitly described in the report, and many of the relationships entered into the SAA tool by participants described only a single factor rather than a relationship between factors. Given that the led outdoor activity sector has only recently been exposed to systems thinking models of accident causation (e.g. Salmon et al, 2010), it is likely that some participants' performance with the tool and SAA method was impacted by a limited understanding of systems thinking, accident causation, and particularly the notion of relationships between different contributory factors. On reason for this may that until the introduction of UPLOADS, the understanding of accident causation in this domain was focused on identifying errors and failures within the immediate context of the incident (Goode et al., 2015). A limited understanding of systems thinking accident models has been previously been implicated in the research-practice gap in accident analysis in other domains (e.g. Salmon, McClure, & Stanton, 2012; Underwood & Waterson, 2013). Moreover, experiences of applying the systems approach to domains such as patient safety, which have considerably greater resources, suggest that translating the systems approach into practice is likely to take a significant amount of time (Waterson, 2009).

In the context of the current study, participant's lack of systems thinking may indicate frailties in the training material. The study did not evaluate the extent to which practitioners utilised this material, so it is unclear whether the material simply was not read or whether it was difficult to understand. While face-to-face training would overcome these issues, it is not practical as the intended end users of UPLOADS are widely distributed across Australia. An additional factor may be high levels of staff turnover within participating organisations. While it is unknown whether this played a role in the current trial, it was identified as a key challenge for implementation of the system in a previous trial (Goode, Salmon, Lenne, & Finch, In Press). Following this, a decision was made to embed the links to the training within the software tool; however, the findings seem to indicate that this had little impact. Insufficient training and limitations on time available have been shown to be a significant problem in other studies evaluating the application of SAA methods by practitioners (Underwood et al., 2016). Further research is required to determine how best to train practitioners to use SAA methods with limited resources.

In terms of the classification of factors and relationships, the study revealed a number of areas for improving the SAA method and how it is implemented within the software tool. First, the results revealed some systematic differences in the way that participants and researchers used particular codes from the taxonomy. The codes identified as problematic in the current study were the same as those identified in a recent reliability study using artificial incident coding tasks (Taylor et al., 2015). This suggests that the definitions provided for these codes are counter-intuitive to practitioners, and should be revised. Second, many participants consistently selected the same factor twice to classify a relationship. This may indicate that that there may be a technical problem with this aspect of the software tool, which should be investigated in future usability studies.

While overall these findings are disappointing, this trial represents an important step towards supporting a deeper understanding of accident causation in the led outdoor activity sector. Prior to the development of UPLOADS, a survey of Australian outdoor activity providers found that only half had incident databases (Goode, Finch, Cassell, Lenne, & Salmon, 2014). Thus, at the very least, organisations that use the UPLOADS Software Tool will benefit by being able to record incidents over time. Secondly, other incident reporting systems for this domain either do not support the identification of contributing factors (Goode, Finch, et al., 2014), or include taxonomies that are limited to factors relating to activity leaders, participants, equipment and the environment (Salmon et al., 2014). Therefore, while users of the UPLOADS Software Tool may not have a complete picture of accident causation in their organisation, they potentially have an enhanced understanding compared to not using an incident database at all, or using any of the other incident reporting systems that have been developed for the domain. To this end, continued use of the tool is appropriate; however, further development is required to ensure that the potential benefits of the tool are realised.

Overall the findings highlight the issues associated with generalising reliability and validity findings from laboratory-based studies to situations where practitioners analyse their own incident data. Experimental control is required to establish that the (positive or negative) findings are attributable to the accident analysis method. However, the findings from this study indicate that does not necessarily ensure that the method is "reliable" or "valid" when it is implemented within a tool for practitioners. In this study, participant's identification and classification of contributing factors was far poorer that could have been predicted from previous studies that utilised a series of incident coding tasks to assess reliability (Taylor et al., 2015). This is problematic as the majority of published reliability and validity studies have used a similar methodology (Olsen, 2013). Many authors recognise that “there is a need for further validation…using more realistic accident scenario exercises” of accident analysis methods, following laboratory-based studies (p.169, Gordon, Flin, & Mearns, 2005). However, the available literature indicates that such follow-up studies are either rarely conducted or suffer from the file drawer effect (Rosenthal, 1979). In line with a recent study conducted by Underwood et al. (2016), this study highlights the importance of evaluating accident analysis methods in the context of the intended use to identify barriers to translation.

Finally, the limitations of the study and directions for future research should be considered. First, participants with few incidents to report had less opportunity to interact with the Software Tool than those who had more incidents. While this was a consequence of the naturalistic study design, it would be useful to evaluate the validity of the software tool under conditions where all participants had to analyse the same number of incident reports. Second, the study design was not ideal for evaluating whether participants were more or less likely than researchers to identify certain types of factors or use certain types of codes from the taxonomy. Again, this was a consequence of using incident reports collected by participants, as we had no control over the content. Third, participants in this study used to the tool for a very limited period of time, and training was self-directed (i.e. participants could choose which of the training materials they viewed). This may partially explain why so few organisations used the SAA tools. The critical question for the next stage of the research is how the training can be integrated into the software tool, so that the application of the accident analysis method becomes intuitive for end users. Fourth, the participants were a self-selecting sample; potentially different results would have been obtained with a broader sample. Follow-up interviews from the trial will investigate why some participants only use the software to collect incident reports, and do not undertake any analysis. Another area that requires further investigation are the factors that influence whether organisations choose to participate in the trial. This is important as all organisation involved in the trial contribute data to a national dataset, that is regularly analysed to inform the sector of the risks it faces.

In conclusion, this study represents a small step towards the translation of SAA into one area of practice, the led outdoor activity domain. Although far from perfect, the UPLOADS Software Tool represents an attempt to ensure that an appropriate model of accident causation underpins incident reporting and analysis by practitioners. More work is required that targets the usability of the tool and practitioner’s understanding of SAA. This work needs to be undertaken in tandem to ensure that the use of the tool ultimately results the identification of more appropriate and effective countermeasures to enhance accident prevention in this domain. From a methodological perspective, the study highlights some of the benefits and difficulties of evaluating the validity of human factors methods using real life incident reports.

Acknowledgements

This project is supported by the Australian Research Council (ARC) in partnership with Australian Camps Association, Outdoor Educators’ Association of South Australia, Outdoors South Australia, United Church Camping, Outdoors Victoria, Outdoor Council of Australia, Recreation South Australia, Outdoor Recreation Industry Council, Outdoors WA, YMCA

Victoria, The Outdoor Education Group, Girl Guides Australia, Wilderness Escape Outdoor Adventures, Venture Corporate Recharge, Queensland Outdoor Recreation Federation, Christian Venues Association, Parks Victoria, Victoria Department of Planning and Community Development, Outdoor Education Australia and the Department of National Parks, Recreation, Sport and Racing Australia (LP110100037). Paul Salmon's contribution was funded through his ARC Future Fellowship (FT140100681). Caroline Finch was supported by a NHMRC Principal Research Fellowship (ID: 565900). The Australian Centre for Research into Injury in Sport and its Prevention (ACRISP) is one of the International Research Centres for Prevention of Injury and Protection of Athlete Health supported by the International Olympic Committee (IOC). Paul Salmon's contribution was funded through his ARC Future Fellowship (FT140100681).

References

Brackenberg, M. (1999). Learning from our mistakes–before it's too late. *Australian Journal of Outdoor Education, 3*(27 - 33).

Branford, K. (2011). Seeing the Big Picture of Mishaps: Applying the AcciMap Approach to Analyze System Accidents *Aviation Psychology and Applied Human Factors, 1*(1).

Cassano-Piche, A.L., Vicente, K. J., & Jamieson, G. A. (2009). A test of Rasmussen's risk management framework in the food safety domain: BSE in the UK. *Theoretical Issues in Ergonomics Science, 10*(4), 283-304. doi: 10.1080/14639220802059232

Cornelissen, Miranda, McClure, Roderick, Salmon, Paul M., & Stanton, Neville A. (2014). Validating the Strategies Analysis Diagram: Assessing the reliability and validity of a formative method. *Applied Ergonomics, 45*(6), 1484-1494.

Curtis, R. (1995). OA Guide to Outdoor Safety Management: Outdoor Action Program, Princeton University.

Goode, N., Finch, C., Cassell, E., Lenne, M.G., & Salmon, P.M. (2014). What would you like? Identifying the required characteristics of an industry-wide incident reporting and learning system for the led outdoor activity sector. *Australian Journal of Outdoor Education, 17*(2).

Goode, N., Salmon, P.M., Lenne, M. & Finch, C.F. (2014). A test of a systems theory-based incident coding taxonomy for risk managers. In Ahram, T., Karwowski, W., Marek, T. (Eds.) *Proceedings of the 5th Applied Human Factors and Ergonomics International Conference* (pp. 5098-5108), The Printing House Inc.

Goode, N., Salmon, P.M., Finch, C.F. & Lenne (2015). Looking beyond people, equipment and environment: Is a systems theory model of accident causation required to understand injuries and near misses during outdoor activities? AHFE2015, 1st International Conference on Human Factors in Sports and Outdoor Recreation. *AHFE 2015 Conference Proceedings* ISBN: 978-1-4951-6042-4

Goode, N., Salmon, P.M., Lenne, M.G., & Finch, C.F. (In Press). The UPLOADS Project: Challenges in implementing an Australia-wide incident reporting system for led outdoor activity providers. *Wilderness and Environmental Medicine*.

Gordon, R., Flin, R., & Mearns, K. (2005). Designing and evaluating a human factors investigation tool (HFIT) for accident analysis. *Safety Science, 43*(3), 147-171.

Hill, A. (2011). National Incident Database Report 2010: Outdoor education and recreation: New Zealand Mountain Safety Council.

Johnson, C.W., & Muniz de Almeida, I. (2008). Extending the borders of accident investigation: applying novel analysis techniques to the loss of the Brazilian space launch vehicle VLS-1 V03. *Safety Science, 46*(1), 38-53.

Leveson, Nancy G. (2011). Applying systems thinking to analyze and learn from events. *Safety Science, 49*(1), 55-64.

Lenné, M. G., Ashby, K. & Fitzharris, M. (2008). Analysis of general aviation crashes in Australia using the Human Factors Analysis and Classification System. International Journal of Aviation Psychology, *18*, 340-352

Lundberg, J., Rollenhagen, C., & Hollnagel, E. (2009). What-You-Look-For-Is-What-You-Find – The consequences of underlying accident models in eight accident investigation manuals. *Safety Science, 47*(10), 1297-1311.

Newnam, S., & Goode, N. (2015). Do not blame the driver: A systems analysis of the causes of road freight crashes *Accident analysis and prevention, 76*, 141-151.

Olsen, N.S. (2013). Reliability studies of incident coding systems in high hazard industries: A narrative review of study methodology. *Applied Ergonomics, 44*(2), 175-184.

Priest, S. & Gass, M.A. (2005). Effective leadership in adventure programming: Human Kinetics.

Rasmussen, J. (1997). Risk management in a dynamic society: A modelling problem. *Safety Science, 27*(2/3), 183-213.

Read, G.J.M., Salmon, P.M., & Lenné, M.G. (2013). Sounding the warning bells: The need for a systems approach to understanding behaviour at rail level crossings. *Applied Ergonomics*, 44, 5, 764-774.

Rosenthal, R. (1979). The "File Drawer Problem" and the Tolerance for Null Results. *Psychological Bulletin, 86*(3), 638-641.

Salmon, P.M, Williamson, A., Lenné, M., Mitsopoulos-Rubens, E., & Rudin-Brown, C.M. (2010). Systems-based accident analysis in the led outdoor activity domain: application and evaluation of a risk management framework. *Ergonomics, 53*(8), 927-939.

Salmon, P.M., Goode, N., Archer, F., Spencer, C., McArdle, D., & McClure, R.J. (2014). A systems approach to examining disaster response: Using Accimap to describe the factors influencing bushfire response. *Safety Science, 70*, 114-122.

Salmon, P.M., Goode, N., Lenné, M.G., Finch, C.F., & Cassell, E. (2014). Injury causation in the great outdoors: A systems analysis of led outdoor activity injury incidents. *Accident Analysis & Prevention, 63*, 111-120.

Salmon, P.M., Goode, N., Taylor, N.Z., Lenne, M.G., Dallat, C., & Finch, C.F. (In Press). Rasmussen's legacy in the great outdoors: a new incident reporting and learning system for led outdoor activities. *Applied Ergonomics*.

Salmon, P.M., Cornelissen, M., & Trotter, M.J. (2012). Systems-based accident analysis methods: A comparison of Accimap, HFACS, and STAMP. *Safety Science, 50*(4), 1158-1170.

Salmon, P.M., McClure, Rod, & Stanton, Neville A. (2012). Road transport in drift? Applying contemporary systems thinking to road safety. *Safety Science, 50*(9), 1829-1838.

Salmon, P.M., Read, G., Stanton, N.A., & Lenné, M. (2013). The Crash at Kerang: Investigating systemic and psychological factors leading to unintentional non-compliance at rail level crossings. *Accident Analysis & Prevention, 50*(1278-1288).

Salmon, P.M., Stanton, N.A., Lenné, M., Jenkins, D.P., Rafferty, L., & Walker, G.H. (2011). *Human Factors Methods and Accident Analysis Practical Guidance and Case Study Applications*. U.K.: Ashgate.

Salmon, P. M., Walker, G. H., Stanton, N. A., Goode, N., Read, G. (In Press). Fitting methods to paradigms: are ergonomics methods fit for systems thinking? *Ergonomics*, Accepted for publication 8th September 2015

Stanton, N.A., & Young, M.S. (1999). What price ergonomics? *Nature, 399*(6733), 197-198.

Stanton, N.A., & Young, M.S. (2003). Giving ergonomics away? The application of ergonomics methods by novices. *Applied Ergonomics, 34*(5), 479-490.

Taylor, N.Z., Goode, N., Salmon, P.M., Lenne, M.G., & Finch, C.F. (2015). *Which code is it? Inter-rater reliability of systems theory-based causal factor taxonomy for the outdoor* *sector*. Proceedings of the 19th Triennial Congress of the International Ergonomics Association, Melbourne, Australia.

Underwood, P., & Waterson, P. (2014). Systems thinking, the Swiss Cheese Model and accident analysis: A comparative systemic analysis of the Grayrigg train derailment using the ATSB, AcciMap and STAMP models. *Accident analysis and prevention, 68*, 75-94.

Underwood, P. & Waterson, P. (2013). Systemic accident analysis: Examining the gap between research and practice. *Accident Analysis & Prevention, 55*, 154-164.

Underwood, P., Waterson, P., & Braithwaite, G. (2016). 'Accident investigation in the wild' – A small-scale, field-based evaluation of the STAMP method for accident analysis. *Safety Science, 82*, 129-143.

Waterson, Patrick. (2009). A critical review of the systems approach within patient safety research. *Ergonomics, 52*(10), 1185-1195.